Journal of Computational and Applied Mathematics 234 (2010) 297–304

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Analytical existence of solutions to a system of nonlinear equations with application

Assi N'Guessan *

Laboratoire Paul Painlevé UMR CNRS 8524 and Ecole Polytechnique Universitaire de Lille Université de Lille 1, 59655 Villeneuve d'Ascq Cedex, France

A R T I C L E I N F O

Article history:
Received 9 September 2008

Keywords:
Existence
Nonlinear system
Schur complement
Newton–Raphson
Road safety measure
Multinomial distribution
Crash data

A B S T R A C T

We study the existence of analytical solutions to a system of nonlinear equations under constraints linked to the analysis of a road safety measure without computing second derivatives. We formally demonstrate this existence of solutions by applying a matrix inversion principle through Schur complement to a subsystem of equations derived from the main system. The analytical results thus obtained are used to construct a simple iterative procedure to look for optimal solutions as well as an initial solution adapted to data of each case study. We illustrate our results with simulated accident data obtained from the setting-up of a road safety measure. The numerical solutions thus obtained are then compared to those given through a classic Newton–Raphson type approach directly applied to the main system.

1. Introduction

In most applied mathematics problems (statistics, probability, numerical analysis, etc.) we are faced with the estimation of unknown parameters which are often functionally dependent. A reliable estimation of these parameters will lead to a reliable mathematical analysis of the studied problem. A lot of approximation methods for solutions (estimation of unknown parameters) need the use of the solution to linear or nonlinear equations, with or without constraints. The most popular and the most frequently used approximation method is still the Newton–Raphson method. Within the framework of multivariate statistics and more particularly in the field of the multidimensional assessment of a road safety measure or of road layout works (setting-up of round-abouts, islands, changes in road markings, etc.) the most frequently used approximation method is the maximum likelihood technique, which is an adaptation of the Newton–Raphson method (see for example [1,2]). This approach of the maximum likelihood under constraints (see for example [3–6]) has recently been used by some authors (see for example [7,8]) to estimate the parameters linked to some statistical models used in the evaluation of a road safety measure applied simultaneously to several test areas with different accident types. The simulated numerical results obtained by these authors clearly show that the approximation method through the technique of maximum likelihood under constraints gives converging numerical solutions. However, we can note, for small numbers of simulated accidents, that the solution vector of their equation system remains relatively far from the true solution vector. Moreover, the numerical procedures set up by these authors are rather complex and are not easy to use for nonspecialists. Finally, the initial solution, which is necessary to start the iterative procedures, is not automated and it is manual. In this paper, we focus on a particular version of the nonlinear equation system under constraints constructed by N'Guessan et al. [7] and which has recently been used to analyse the impact of the road layout works. More precisely, we demonstrate the existence of analytical solutions to their nonlinear equation system under constraints by applying the Schur complement inversion matrix principle to a subsystem of nonlinear equations. Then we use the results to construct an initial solution adapted to data of each case

* Corresponding address: Bât. Polytech'Lille, Université de Lille 1, 59655 Villeneuve d'Ascq, France. Fax: +33 3 28 76 73 01.
E-mail addresses: Assi.Nguessan@polytech-lille.fr, assi.aguessan@polytech-lille.fr.

0377-0427/$ – see front matter
doi:10.1016/j.cam.2009.12.026

study as well as a simple iterative procedure to look for the optimal solution. Section 2 poses the problem dealt with in this paper along with the notations used after. Section 3 presents and demonstrates the whole of the technical results. Four fundamental lemmas are given and the technical details of their demonstration are given in the Appendix. The main results are given in Section 3.2. Those results are given under the form of two theorems. In Section 4, we illustrate our approximation method thanks to simulated accident data coming from the setting-up of a road safety measure. The obtained optimal solutions are compared to those given with a Newton–Raphson approach directly applied to the same equation system. On average, the numerical convergence is much quicker when we use our results than those from Newton–Raphson. Moreover, our results enable to avoid some fairly complicated intermediate calculations that the Newton–Raphson approach requires.

2. Description of the problem

Let's note R ($R > 1$) a given natural integer, $\mathbf{X}_1 = (x_{11}, x_{12}, \ldots, x_{1R})^T$, $\mathbf{X}_2 = (x_{21}, x_{22}, \ldots, x_{2R})^T$ and $\mathbf{Z} = (c_1, c_2, \ldots, c_R)^T$ three $R \times 1$ vectors of numerical data such that $x_{tj} \geq 0$, $c_j > 0$ with ($t = 1, 2; j = 1, 2, \ldots, R$), and consider $\mathbf{\Theta} = (\beta_0, p_1, p_2, \ldots, p_R)^T$ a $(1+R) \times 1$ parameter vector of unknown components such that $\beta_0 > 0$ and

$$0 < \hat{p}_j \quad \text{with} \sum_{j=1}^{R} \hat{p}_j = 1. \tag{1}$$

The three vectors $\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{Z}$ being supposedly known, the main aim of this paper is to look for an analytical solution, if any, written $\hat{\mathbf{\Theta}} = (\hat{\beta}_0, \hat{p}_1, \hat{p}_2, \ldots, \hat{p}_R)^T$, for the nonlinear equation system with respect to vector $\mathbf{\Theta}$:

$$\begin{cases} \displaystyle\sum_{j=1}^{R} \left\{ x_{2j} - x_{\bullet j} \frac{\hat{\beta}_0 \sum_{m=1}^{R} c_m \hat{p}_m}{1 + \hat{\beta}_0 \sum_{m=1}^{R} c_m \hat{p}_m} \right\} = 0 \\ x_{\bullet j} - n \dfrac{\hat{p}_j (1 + c_j \hat{\beta}_0)}{1 + \hat{\beta}_0 \sum_{m=1}^{R} c_m \hat{p}_m} = 0, \quad (j = 1, 2, \ldots, R); \quad 0 < \hat{p}_j \quad \text{with} \displaystyle\sum_{j=1}^{R} \hat{p}_j = 1 \end{cases} \tag{2}$$

with $x_{\bullet j} = x_{1j} + x_{2j}$, $n = \sum_{j=1}^{R} x_{\bullet j}$. Different iteratives methods (see for instance Orthega [2], Fletcher [1]) can be used to look for a solution to system (2). The Newton–Raphson method for the calculation of a solution is probably the most widely used iterative method. It is well known that this method is quadratically convergent because it needs the computation of second derivatives. We present here a straightforward approach concerning the existing of the analytical solutions to system (2) without computing second derivatives.

With the purpose of looking for a solution to (2), we set the first component β_0 and solve the second subsystem in relation to the components $p_1, p_2, \ldots, p_R$ using Schur complement matrix inversion principle. Then we use the solution to solve the first equation of (2) with respect to β_0, and vice versa. In the following, we note $\mathbf{X}_\bullet = \mathbf{X}_1 + \mathbf{X}_2$, $\mathbf{1}_R = (1, \ldots, 1)^T$, two vectors of dimension R and M_{β_0}, the $R \times R$ diagonal matrix:

$$M_{\beta_0} = \mathrm{diag}\,(1 + \beta_0 c_1, \ldots, 1 + \beta_0 c_R)\,. \tag{3}$$

We then set

$$M_{\beta_0}^{(0)} = \begin{bmatrix} M_{\beta_0} & \dfrac{\beta_0^{1/2}}{n^{1/2}} \mathbf{X}_\bullet \\ \dfrac{\beta_0^{1/2}}{n^{1/2}} \mathbf{Z}^T & 0 \end{bmatrix}, \qquad M_{\beta_0}^{(1)} = \begin{bmatrix} M_{\beta_0} & \dfrac{\beta_0^{1/2}}{n^{1/2}} \mathbf{X}_\bullet \\ \dfrac{\beta_0^{1/2}}{n^{1/2}} \mathbf{Z}^T & 1 \end{bmatrix}, \qquad M_{\beta_0}^{(2)} = \begin{bmatrix} M_{\beta_0} & \dfrac{1}{n^{1/2}} \mathbf{X}_\bullet \\ \dfrac{1}{n^{1/2}} \mathbf{1}_R^T & 0 \end{bmatrix},$$

three $(1+R) \times (1+R)$ matrices thus defined.

3. Fundamental results concerning the existence

We present here the main results concerning the existence of the analytical solutions to system (2). These fundamental results are given through Theorems 3.5 and 3.6 hereafter. These main results require some intermediate results we give in the following technical lemmas. The proof of lemmas are given in the Appendix.

3.1. Preliminary results

Lemma 3.1. *For all set value of β_0, the Schur complement, noted $(M_{\beta_0}^{(1)}/M_{\beta_0})$, of M_{β_0} in $M_{\beta_0}^{(1)}$ exists, and furthermore we have: $(M_{\beta_0}^{(1)}/M_{\beta_0}) > 0$.*

Lemma 3.2. *$(M^{(0)}_{\beta_0}/M_{\beta_0})$ and $(M^{(2)}_{\beta_0}/M_{\beta_0})$ respectively Schur complement of M_{β_0} in $M^{(0)}_{\beta_0}$ (resp. in $M^{(2)}_{\beta_0}$) exist and we have:*

$$(M^{(0)}_{\beta_0}/M_{\beta_0}) + (M^{(2)}_{\beta_0}/M_{\beta_0}) = -1.$$

Lemma 3.3. *$(M^{(1)}_{\beta_0}/1)$ the Schur complement of 1 in $M^{(1)}_{\beta_0}$ exists, it is nonsingular and we have the following two results:*

(i) $(M^{(1)}_{\beta_0}/1)^{-1} = M^{-1}_{\beta_0} + \frac{\beta_0}{n} M^{-1}_{\beta_0} \mathbf{X}_{\bullet} (M^{(1)}_{\beta_0}/M_{\beta_0})^{-1} \mathbf{Z}^T M^{-1}_{\beta_0}$;

(ii) $\frac{1}{n} \mathbf{1}^T_R (M^{(1)}_{\beta_0}/1)^{-1} \mathbf{X}_{\bullet} - 1 = 0$.

Lemma 3.4. *Let's set*

$$\Omega_{\beta_0} = \begin{bmatrix} 1 + \left(1 - \frac{x_{\bullet 1}}{n}\right)\beta_0 c_1 & -\frac{x_{\bullet 1}}{n}\beta_0 c_2 & \cdots & -\frac{x_{\bullet 1}}{n}\beta_0 c_R \\ -\frac{x_{\bullet 2}}{n}\beta_0 c_1 & 1 + \left(1 - \frac{x_{\bullet 2}}{n}\right)\beta_0 c_2 & \cdots & 1 - \frac{x_{\bullet 2}}{n}\beta_0 c_R \\ \vdots & \ddots & \ddots & \vdots \\ -\frac{x_{\bullet R}}{n}\beta_0 c_1 & -\frac{x_{\bullet R}}{n}\beta_0 c_2 & \cdots & 1 + \left(1 - \frac{x_{\bullet R}}{n}\right)\beta_0 c_R \end{bmatrix}.$$

$R \times R$ matrix thus defined. Then, the $(1 + R) \times (1 + R)$ matrix

$$\Sigma_{\beta_0} = \begin{bmatrix} \Omega_{\beta_0} & \mathbf{1}_R \\ \mathbf{1}^T_R & 0 \end{bmatrix}$$

is nonsingular and

$$\Sigma^{-1}_{\beta_0} = \begin{bmatrix} W_{\beta_0} & \|\mathbf{1}_R\|^{-2}_{(M^{(1)}_{\beta_0}/1)^{-1}} (M^{(1)}_{\beta_0}/1)^{-1}\mathbf{1}_R \\ \|\mathbf{1}_R\|^{-2}_{(M^{(1)}_{\beta_0}/1)^{-1}} \mathbf{1}^T_R (M^{(1)}_{\beta_0}/1)^{-1} & -\|\mathbf{1}_R\|^{-2}_{(M^{(1)}_{\beta_0}/1)^{-1}} \end{bmatrix}$$

where

$$W_{\beta_0} = (M^{(1)}_{\beta_0}/1)^{-1} - \|\mathbf{1}_R\|^{-2}_{(M^{(1)}_{\beta_0}/1)^{-1}} (M^{(1)}_{\beta_0}/1)^{-1} \mathbf{1}_R \mathbf{1}^T_R (M^{(1)}_{\beta_0}/1)^{-1}$$

is the $R \times R$ thus defined and

$$\|\mathbf{1}_R\|^2_{(M^{(1)}_{\beta_0}/1)^{-1}} = \mathbf{1}^T_R \times (M^{(1)}_{\beta_0}/1)^{-1} \times \mathbf{1}_R$$

is the square of the norm of vector $\mathbf{1}_R$ with respect to the matrix $(M^{(1)}_{\beta_0}/1)^{-1}$.

3.2. Existence of analytical solutions

Theorem 3.5. *For all set β_0 ($\beta_0 > 0$), the nonlinear subsystem with respect to the components of unknown vector* **P**

$$x_{\bullet j} - n \frac{p_j(1 + c_j\beta_0)}{1 + \beta_0 \sum\limits_{m=1}^{R} c_m p_m} = 0, \quad (j = 1, 2, \ldots, R); \quad 0 < p_j \quad \textit{with} \sum_{j=1}^{R} p_j = 1. \tag{4}$$

accepts as a solution vector $\mathbf{P}_n(\beta_0)$ of dimension R given by

$$\mathbf{P}_n(\beta_0) = \frac{1}{n}(M^{(1)}_{\beta_0}/M_{\beta_0})^{-1} \left[\frac{x_{\bullet 1}}{1 + \beta_0 c_1}, \frac{x_{\bullet 2}}{1 + \beta_0 c_2}, \ldots, \frac{x_{\bullet R}}{1 + \beta_0 c_R}\right]^T. \tag{5}$$

Proof. We show, after a few transformations, that the system (4) above is equivalent to the following linear system:

$$\begin{bmatrix} \Omega_{\beta_0} & \mathbf{1}_R \\ \mathbf{1}^T_R & 0 \end{bmatrix} \times \begin{bmatrix} \mathbf{P} \\ 0 \end{bmatrix} = \begin{bmatrix} n^{-1}\mathbf{X}_{\bullet} \\ 1 \end{bmatrix}, \tag{6}$$

where $\mathbf{P} = (p_1, p_2, \ldots, p_R)^T$ and $n^{-1}\mathbf{X}_{\bullet} = (\frac{x_{\bullet 1}}{n}, \frac{x_{\bullet 2}}{n}, \ldots, \frac{x_{\bullet R}}{n})^T$ are two $R \times 1$ vectors thus defined and Ω_{β_0} the $R \times R$ matrix given in Lemma 3.4. From the expression of $\Sigma^{-1}_{\beta_0}$ of Lemma 3.4, we then deduce, for all set $\beta_0 > 0$, that the solution $\mathbf{P}_n(\beta_0)$

of system (4) is given by

$$\begin{aligned}\mathbf{P}_n(\beta_0) &= W_{\beta_0} \times \frac{\mathbf{X}_\bullet}{n} + \|\mathbf{1}_R\|^{-2}_{(M^{(1)}_{\beta_0}/1)^{-1}} \times (M^{(1)}_{\beta_0}/1)^{-1} \times \mathbf{1}_R \\ &= (M^{(1)}_{\beta_0}/1)^{-1}\frac{\mathbf{X}_\bullet}{n} - \|\mathbf{1}_R\|^{-2}_{(M^{(1)}_{\beta_0}/1)^{-1}}(M^{(1)}_{\beta_0}/1)^{-1}\mathbf{1}_R\left[\frac{1}{n}\mathbf{1}_R^T(M^{(1)}_{\beta_0}/1)^{-1}\mathbf{X}_\bullet - 1\right].\end{aligned}$$

Using now Lemma 3.3 and after a few transformations, we obtain:

$$\begin{aligned}\mathbf{P}_n(\beta_0) &= (M^{(1)}_{\beta_0}/1)^{-1}\frac{\mathbf{X}_\bullet}{n} \\ &= M^{-1}_{\beta_0}\frac{\mathbf{X}_\bullet}{n} + (M^{(1)}_{\beta_0}/M_{\beta_0})^{-1} \times \left(\frac{\beta_0}{n}\mathbf{Z}^T M^{-1}_{\beta_0}\mathbf{X}_\bullet\right) \times M^{-1}_{\beta_0}\frac{\mathbf{X}_\bullet}{n} \\ &= \left[1 + (M^{(1)}_{\beta_0}/M_{\beta_0})^{-1}(1 - (M^{(1)}_{\beta_0}/M_{\beta_0}))\right] M^{-1}_{\beta_0}\frac{\mathbf{X}_\bullet}{n} \\ &= (M^{(1)}_{\beta_0}/M_{\beta_0})^{-1} M^{-1}_{\beta_0}\frac{\mathbf{X}_\bullet}{n}. \quad \blacksquare\end{aligned}$$

Note. We check that solution vector $\mathbf{P}_n(\beta_0)$ thus obtained really belongs to the order R simplex. Indeed, applying Lemma 3.2, we have:

$$\begin{aligned}\mathbf{1}_R^T \times \mathbf{P}_n(\beta_0) &= \frac{1}{n}(M^{(1)}_{\beta_0}/M_{\beta_0})^{-1}\sum_{j=1}^{R}\frac{x_{\bullet j}}{1+\beta_0 c_j} \\ &= (M^{(1)}_{\beta_0}/M_{\beta_0})^{-1} \times \left(1 - \frac{\beta_0}{n}\mathbf{Z}^T M^{-1}_{\beta_0}\mathbf{X}_\bullet\right) \\ &= 1.\end{aligned}$$

Theorem 3.6. *Using the results of Theorem 3.5, we demonstrate that nonlinear equation system (2) accepts a solution*

$$\hat{\mathbf{\Theta}} = (\hat{\beta}_0, \hat{p}_1, \ldots, \hat{p}_R)^T$$

of components

$$\begin{cases}\hat{\beta}_0 = \dfrac{\sum\limits_{m=1}^{R} x_{2\,m}}{\left(\sum\limits_{m=1}^{R} c_m\hat{p}_m\right) \times \left(\sum\limits_{m=1}^{R} x_{1\,m}\right)}; \\ \hat{p}_j = \dfrac{1}{n}(M^{(1)}_{\hat{\beta}_0}/M_{\hat{\beta}_0})^{-1} \times \dfrac{x_{\bullet j}}{1+\hat{\beta}_0 c_j}, \quad (j = 1, 2, \ldots, R).\end{cases} \tag{7}$$

4. An industrial application

4.1. A statistical road accident data modelling

We apply the above results to the estimation of a road safety measure effect in presence of different accident risks. We consider a multidimensional combination of road accident frequencies before and after the introduction of a road safety measure (crossroad layout, surface of motorway, etc.) at a single experimental site which counts R ($R > 1$) accident types (fatal accidents, accidents with seriously injured people, with slightly injured people etc.) over a fixed period of time. In order to take some external factors into account (such as traffic flow, speed limit variation, weather conditions, ...), the experimental site is matched to a control site where the safety measure has not been directly applied. The $R \times 1$ vector $\mathbf{X}_1$ (resp. $\mathbf{X}_2$) denotes the R accident types number on experimental site before (resp. after) the introduction of the safety measure. The $R \times 1$ vector $\mathbf{Z}$ denotes the vector of accident data for control site over the same time period where z_j denotes the ratio of the number of type j accidents for the period after to the period before.

N'Guessan et al. [7,8] propose different statistical models to combine road accident frequencies when a road safety measure is applied on different sites. Their method of estimating simultaneously the average effect and the accident risks is based on the maximization of the so-called likelihood function using the Newton–Raphson (NR) approach and NAG [9] software of optimisation algorithms. We propose a new approach (NA) based on N'Guessan and Truffier's algorithm [10] of

Table 1
Accident data set.

n	Before			After		
18	7	1	7	2	1	0
42	16	3	14	4	2	3
2141	636	114	621	276	100	394

Table 2
Numerical approximation.

Parameter	*NR*	*NA*
	$n = 18$	
β_0	0.1308	0.1308
p_1	0.5145	0.5144
p_2	0.1009	0.1009
p_3	0.3846	0.3846
	$n = 42$	
β_0	0.1771	0.1889
p_1	0.4947	0.4956
p_2	0.1056	0.1050
p_3	0.3996	0.3993
	$n = 2141$	
β_0	0.3658	0.4910
p_1	0.4541	0.4798
p_2	0.0823	0.0700
p_3	0.4636	0.4502

estimation of the average effect and the accident risks using Theorem 3.6 and a particular case of the model of N'Guessan et al. [7] where the logarithm of their likelihood function used is given by

$$\mathcal{L}(\mathbf{\Theta}) = \text{constant} + \sum_{j=1}^{R} \left\{ x_{\bullet j} \log_e(p_j) + x_{2j} \log_e(\beta_0) - x_{\bullet j} \log_e \left(1 + \beta_0 \sum_{m=1}^{r} c_m p_m \right) \right\} \tag{8}$$

such that vector $\mathbf{P}$ belongs to the simplex $\mathcal{S}^{(R-1)}$:

$$\mathbf{P} \in \mathcal{S}^{(R-1)} = \left\{ (p_1, p_2, \ldots, p_R)^T \in \mathbb{R}^R, p_j > 0, \sum_{j=1}^{R} p_j = 1 \right\}. \tag{9}$$

So the vector of parameters $\mathbf{\Theta}$ has the following restraints $\beta_0 > 0$, $(0 < p_j < 1)$ and the linear constraints

$$h(\mathbf{\Theta}) = 0, \quad \text{with } h(\mathbf{\Theta}) = \langle \mathbf{1}_R, \mathbf{P} \rangle - 1, \tag{10}$$

where $\langle , \rangle$ is the classic inner product. So taking the first derivatives of $\mathcal{L}(\mathbf{\Theta})$ with respect to each component of the vector parameter $\mathbf{\Theta}$, setting them equal to zero and adding the constraints $\beta_0 > 0$, $p_j > 0$ and $\sum_{j=1}^{R} p_j = 1$ (using the Lagrange multiplier), we show that the maximum likelihood equations are equivalent to the system (2).

To find the solutions to system (2), N'Guessan et al. [7] used a FORTRAN program where they proceeded iteratively with the Newton–Raphson type algorithm available in the NAG software [9]. In the present paper, we use Theorem 3.6 and Microsoft Excel program to find iteratively the solutions to system (2). Then we compare the two set of solutions using the same simulated set data. A few numerical simulation results are displayed in the next section.

4.2. *A numerical simulation application*

The main object of the simulation is first to look for a solution to the constrained equations using our approach (NA) and secondly to compare our results to those obtained by the latter authors using NR approach. Without loss of generality, we suppose $R = 3$ (fatal accidents, accidents with seriously injured people, with slightly injured people) and we use the same simulation principle of before–after accident data with control site of N'Guessan et al. [7,8]. Several accident data sets have been generated using different values of n and by setting

$$\mathbf{Z} = (1.2707, 2.4512, 1.6317)^T, \qquad \mathbf{\Theta}^0 = (0.5000, 0.4619, 0.0776, 0.4605)^T$$

where $\mathbf{Z}$ represents the data set of the control site and $\mathbf{\Theta}^0$ represents the true value of the unknown parameter vector $\mathbf{\Theta}$ which has to be estimated. The numerical simulated data (Table 1) and simulated results (Table 2) below are found for n belonging to $\{18, 42, 2141\}$. In order to compare numerically NR approach and NA approach, we use the same initial point

$$\mathbf{\Theta}^{(0)} = (0.2000, 0.3333, 0.3333, 0.3333)^T$$

i.e. the initial point used by the latter authors.

Each line of data in Table 1 represents the repartition of n, the total accident number on the experimental site among the concerned three accident types and between the periods before and after the setting-up of the road safety measure. For example, for $n = 18$, we have $\mathbf{X}_1^T = (7, 1, 7)$ (resp. $\mathbf{X}_2^T = (2, 1, 0)$). This means that the number of fatal accidents observed on the experimental site has decreased from 7 in the period before the setting-up to 2 in the period after. The number of accidents with seriously injured people has remained steady (from 1 to 1) and the number of accidents with light injuries went down from 7 to 0.

For the same periods of time, the ratio of the total number of fatal accidents after to the one before on the control site is 1.2707 (first component of vector $\mathbf{Z}$), i.e. a 27% increase of fatal accidents observed on the control site. In the same periods, the accidents with serious injuries (resp. light injuries) on the control site have increased by about 145% (resp. 63%). Using the whole set of those simulated data (Table 1), the components of vector $\mathbf{Z}$ and the values of vector $\mathbf{\Theta}^{(0)}$ as the initial solution, we get the approximations given in Table 2.

The whole set of the simulation results shows that the solutions to system (2) obtained through the NA approach are similar to those obtained through the NR approach, indeed better in some cases. For example, for fairly high values of n, the best solution vectors to system (2) are obtained through the NA approach. Likewise, we note that the iteration number of the NA approach is much inferior to the one of the NR approach. For the three case studies presented here, the iteration number of the NR approach is about 10 whereas the one for the NA approach varies between 2 and 5. On average, the numerical convergence is much quicker when we use the NA approach. Moreover, this approach enables to avoid some fairly complicated intermediate calculations that the NR approach requires. In particular, with the NA approach, we do not necessarily have, for each iteration, the obligation of the calculation and inversion of the matrix of the second derivatives. Finally, the NA approach enables to automate and give an initial solution adapted to the study data. It is easy to use because it doesn't require complex and difficult to obtain numerical iterative procedures such as the ones used in the NR approach.

5. Concluding remarks

The NA approach can also be generalised to the solution of some equation systems used in the framework of the modelling of a road safety measure applied simultaneously to different experimental sites and several accident types (see [7,8]). We will then have to compare, with the help of the proper criteria, the generalised versions of the NA approach and of the NR approach used by the latter authors. Such a generalisation of the NA approach would enable to replace the simultaneous search for solutions by a search in several blocks where the solution space dimension is much smaller. We consider completing this new iterative approach of solution with the recent results concerning the formal inversion through Schur complement of some matrices used in road safety (see for example [11–14]). Coupling those results which only use the Shur complement approach will then enable the implementation of help in decision making for people in charge of managing roads and motorways.

Acknowledgements

Part of this paper was done while the author was an invited professor at CIRRELT (Interuniversity Research Centre on Entreprise Networks, Logistics and Transportation) of Université de Montréal. He heartily thanks the staff and professor François Bellavance for their support.

Appendix. Proof of Lemmas

Proof of Lemma 3.1. As $\forall\,\beta_0$, the $R \times R$ matrix M_{β_0} is nonsingular then, by definition, the Schur complement of M_{β_0} in $M_{\beta_0}^{(1)}$ exists. Moreover,

$$\begin{aligned}(M_{\beta_0}^{(1)}/M_{\beta_0}) &= 1 - \frac{\beta_0}{n}\mathbf{Z}^T M_{\beta_0}^{-1}\mathbf{X}_\bullet \\ &= 1 - \frac{1}{n}\sum_{j=1}^{R}\frac{x_{\bullet j}\beta_0 c_j}{1+\beta_0 c_j} \\ &> 0,\end{aligned}$$

because $\frac{1}{n}\sum_{j=1}^{R}\frac{x_{\bullet j}\beta_0 c_j}{1+\beta_0 c_j} < 1$. ■

Proof of Lemma 3.2. A few matrix manipulations enable to show that

$$\begin{aligned}(M_{\beta_0}^{(0)}/M_{\beta_0}) + (M_{\beta_0}^{(2)}/M_{\beta_0}) &= -\frac{\beta_0}{n}\mathbf{Z}^T M_{\beta_0}^{-1}\mathbf{X}_\bullet - \frac{1}{n}\mathbf{1}_R^T M_{\beta_0}^{-1}\mathbf{X}_\bullet \\ &= -\frac{1}{n}\sum_{j=1}^{R} x_{\bullet j} \\ &= -1. \quad ■\end{aligned}$$

Proof of Lemma 3.3. (i) By definition $(M_{\beta_0}^{(1)}/1) = M_{\beta_0} - (\beta_0/n)\mathbf{X}_\bullet\mathbf{Z}^T$. Using Lemma 3.1 and the properties linking the inversion of two Schur complements (see for example [15,16]) of a same matrix, we have

$$\begin{aligned}(M_{\beta_0}^{(1)}/1)^{-1} &= (M_{\beta_0} - (\beta_0/n)\mathbf{X}_\bullet\mathbf{Z}^T)^{-1} \\ &= M_{\beta_0}^{-1} + \frac{\beta_0}{n}M_{\beta_0}^{-1}\mathbf{X}_\bullet(M_{\beta_0}^{(1)}/M_{\beta_0})^{-1}\mathbf{Z}^T M_{\beta_0}^{-1}.\end{aligned}$$

(ii) This item's proof uses item (i) of Lemmas 3.3, 3.2 and a few matrix manipulations. We hereafter give the technical details.

$$\begin{aligned}\frac{1}{n}\mathbf{1}_R^T(M_{\beta_0}^{(1)}/1)^{-1}\mathbf{X}_\bullet &= \frac{1}{n}\mathbf{1}_R^T\left[M_{\beta_0}^{-1} + \frac{\beta_0}{n}M_{\beta_0}^{-1}\mathbf{X}_\bullet(M_{\beta_0}^{(1)}/M_{\beta_0})^{-1}\mathbf{Z}^T M_{\beta_0}^{-1}\right]\mathbf{X}_\bullet \\ &= \frac{1}{n}\mathbf{1}_R^T M_{\beta_0}^{-1}\mathbf{X}_\bullet + \frac{1}{n}\left[\frac{\beta_0}{n}\mathbf{1}_R^T M_{\beta_0}^{-1}\mathbf{X}_\bullet(M_{\beta_0}^{(1)}/M_{\beta_0})^{-1}\mathbf{Z}^T M_{\beta_0}^{-1}\mathbf{X}_\bullet\right].\end{aligned}$$

According to Lemma 3.2, we have

$$1 = \frac{\beta_0}{n}\mathbf{Z}^T M_{\beta_0}^{-1}\mathbf{X}_\bullet + \frac{1}{n}\mathbf{1}_R^T M_{\beta_0}^{-1}\mathbf{X}_\bullet.$$

We therefore deduce that

$$\begin{aligned}\frac{1}{n}\mathbf{1}_R^T(M_{\beta_0}^{(1)}/1)^{-1}\mathbf{X}_\bullet &= \frac{1}{n}\mathbf{1}_R^T[M_{\beta_0}^{-1} + \frac{\beta_0}{n}M_{\beta_0}^{-1}\mathbf{X}_\bullet(M_{\beta_0}^{(1)}/M_{\beta_0})^{-1}\mathbf{Z}^T M_{\beta_0}^{-1}]\mathbf{X}_\bullet \\ &= \left(1 - \mathbf{Z}^T M_{\beta_0}^{-1}\mathbf{X}_\bullet\right) \times \left(1 + (M_{\beta_0}^{(1)}/M_{\beta_0})^{-1}\frac{\beta_0}{n}\mathbf{Z}^T M_{\beta_0}^{-1}\mathbf{X}_\bullet\right) \\ &= (M_{\beta_0}^{(1)}/M_{\beta_0}) \times \left(1 + [1 - (M_{\beta_0}^{(1)}/M_{\beta_0})] \times (M_{\beta_0}^{(1)}/M_{\beta_0})^{-1}\right) \\ &= 1. \quad \blacksquare\end{aligned}$$

Proof of Lemma 3.4.

$$\begin{aligned}\Omega_{\beta_0} &= M_{\beta_0} - \frac{\beta_0}{n}\mathbf{X}_\bullet\mathbf{Z}^T \\ &= (M_{\beta_0}^{(1)}/1).\end{aligned}$$

Item (i) of Lemma 3.3 implies that Ω_{β_0} is nonsingular and $\Omega_{\beta_0}^{-1} = (M_{\beta_0}^{(1)}/1)^{-1}$. So $(\Sigma_{\beta_0}/\Omega_{\beta_0})$ the Schur complement of Ω_{β_0} in Σ_{β_0} exists and

$$\begin{aligned}(\Sigma_{\beta_0}/\Omega_{\beta_0}) &= -\mathbf{1}_R^T\Omega_{\beta_0}^{-1}\mathbf{1}_R \\ &= -\|\mathbf{1}_R\|^2_{(M_{\beta_0}^{(1)}/1)^{-1}}.\end{aligned} \tag{11}$$

Using, now, Lemma 3.1 and some properties of the formal inversion of a partitioned matrix through Schur complement (see for instance [15,16]) Σ_{β_0} is nonsingular and

$$\Sigma_{\beta_0}^{-1} = \begin{bmatrix} W_{\beta_0} & -\Omega_{\beta_0}^{-1}\mathbf{1}_R(\Sigma_{\beta_0}/\Omega_{\beta_0})^{-1} \\ -(\Sigma_{\beta_0}/\Omega_{\beta_0})^{-1}\mathbf{1}_R^T\Omega_{\beta_0}^{-1} & (\Sigma_{\beta_0}/\Omega_{\beta_0})^{-1}\end{bmatrix} \tag{12}$$

where

$$\begin{aligned}W_{\beta_0} &= \Omega_{\beta_0}^{-1} + \Omega_{\beta_0}^{-1}\mathbf{1}_R(\Sigma_{\beta_0}/\Omega_{\beta_0})^{-1}\mathbf{1}_R^T\Omega_{\beta_0}^{-1} \\ &= (M_{\beta_0}^{(1)}/1)^{-1} - \|\mathbf{1}_R\|^{-2}_{(M_{\beta_0}^{(1)}/1)^{-1}}(M_{\beta_0}^{(1)}/1)^{-1}\mathbf{1}_R\mathbf{1}_R^T(M_{\beta_0}^{(1)}/1)^{-1}\end{aligned}$$

is the $R \times R$ matrix thus defined and

$$\|\mathbf{1}_R\|^2_{(M_{\beta_0}^{(1)}/1)^{-1}} = \mathbf{1}_R^T \times (M_{\beta_0}^{(1)}/1)^{-1} \times \mathbf{1}_R$$

is the square of the norm of vector $\mathbf{1}_R$ with respect to the matrix $(M_{\beta_0}^{(1)}/1)^{-1}$. $\blacksquare$

References

[1] R. Fletcher, Practical Methods of Optimization: Constrained Optimization, Vol. 2, John W. Chichester, 1981.
[2] J.M. Ortega, W.C. Rheinboldt, Iterative Solutions of Nonlinear Equations in Several Variables, Academic, New York, 1970.

[3] J. Aitchison, S.D. Silvey, Maximum likelihood estimation of parameters subject to restraints, Ann. Math. Stat. 29 (1958) 813–829.
[4] M. Crowder, On the constrained maximum likelihood estimation with non i.i.d. observations, Ann. Inst. Statist. Math. 36A (1984) 239–249.
[5] M. Haber, M.B. Brown, Maximum likelihood methods for log-linear models when expected frequencies are subject to linear constraints, J. Amer. Statist. Assoc. 81 (394) (1986) 477–482.
[6] G.B. Matthews, N.A.S. Crowther, A maximum likelihood estimation procedure when modelling in terms of constraints, South African Statist. J. (29) (1995) 29–50.
[7] A. N'Guessan, A. Essai, C. Langrand, Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière, Rev. Statistique Appliquée XLIX (2) (2001) 83–100.
[8] A. N'Guessan, A. Essai, M. N'Zi, An estimation method of the average effect and the different accident risks when modelling a road safety measure: A simulation study, Comput. Stat. Data Anal. 51 (2006) 1260–1277.
[9] NAG, Numerical Algorithms Groups Limited Ecole Polytechnique Universitaire de Lille, Université des Sciences et Technologies de LILLE, 2003.
[10] A. N'Guessan, M. Truffier, Impact d'un aménagement de sécurité routière sur la gravité des accidents de la route, Journal de la Société Française de Statistique 149 (2) (2008) 99–117.
[11] A. N'Guessan, Constrained covariance matrix estimation in road accident modelling with Shur complements, C.R. Acad. Sci. Paris, Ser. I 337 (2003) 219–222.
[12] A. N'Guessan, C. Langrand, A covariance components estimation procedure when modelling a road safety measure in terms of linear constraints, Statistics 39 (4) (2005) 303–314.
[13] A. N'Guessan, C. Langrand, A Schur complement approach for computing subcovariance matrices arising in a road safety measure modelling, J. Comput. Appl. Math. 177 (2005) 331–345.
[14] A. N'Guessan, F. Bellavance, A confidence interval estimation problem using the Schur complement approach, with application, C.R. Math. Rep. Acad. Sci. Canada 27 (3) (2005) 84–91.
[15] D.V. Ouellette, Schur complements and statistics, Linear Algebra Appl. 36 (1981) 187–295.
[16] F. Zhang (Ed.), Schur Complement and Its Applications, Springer Verlag, 2005.